Published in final edited form as:

Am Heart J. 2016 February ; 172: 34–41. doi:10.1016/j.ahj.2015.10.009.

Evaluation of a Provocative Dyspnea Severity Score in Acute Heart Failure

Omar F. AbouEzzeddine, MDCM¹, Anuradha Lala, MD^{2,3}, Prateeti P. Khazanie, MD⁴, Ravi Shah, MD⁵, Jennifer E. Ho, MD⁶, Horng H. Chen, MBBCh¹, Peter S. Pang, MD⁷, Steven E. McNulty, MS⁸, Kevin J. Anstrom, PhD⁸, Adrian F. Hernandez, MD, MHS⁸, Margaret M. Redfield, MD¹, and for the NHLBI Heart Failure Clinical Research Network

¹Mayo Clinic, Rochester, MN

²Brigham and Women's Hospital, Boston, MA

³NYU Langone Medical Center, New York University School of Medicine, New York, NY

⁴Duke University Medical Center and Duke Heart Center, Durham, NC

⁵Massachusetts General Hospital, Boston, MA

⁶Boston University School of Medicine, Boston, MA

⁷Indiana University School of Medicine, Indianapolis, Indiana

⁸Duke Clinical Research Institute, Durham, NC

Abstract

Background—The acute heart failure (AHF) Syndromes International Working Group proposed that dyspnea be assessed under standardized, incrementally provocative maneuvers and called for studies to assess the feasibility of this approach. We sought to assess the feasibility and statistical characteristics of a novel provocative dyspnea severity score (pDS) versus the traditional dyspnea visual analogue scale (DVAS) in an AHF trial.

Methods—At enrollment, 24, 48 and 72 hours, 230 ROSE-AHF patients completed a DVAS. Dyspnea was then assessed with five-point Likert dyspnea scales administered during four stages (A: upright-with O₂, B: upright-without O₂, C: supine-without O₂ and D: exercise-without O₂). Patients with moderate or less dyspnea were eligible for the next stage.

Correspondence: Omar F. AbouEzzeddine, MDCM; Mayo Clinic, 200 First Street SW Rochester, MN 55905; Abouezzeddine.Omar@mayo.edu; Fax: 507-255-2550; Phone: 507-284-3545.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

HHC has no specific COI related to this work; however, he reports that he and Mayo Clinic have patented and licensed designer natriuretic peptides to Anexon Inc and Capricor Therapeutics. **PSP** has no specific COI related to this work; however, he has or has had the following relationships with the following entities over the last 3 years where a conflict of interest might be perceived given dyspnea is a common endpoint in clinical trials. **AFH** has no specific COI related to this work.

DISCLOSURES

All others report no disclosures relevant to this manuscript.

Results—At enrollment, oxygen withdrawal and supine provocation were highly feasible (97%), provoking more severe dyspnea (1 Likert point) in 24% and 42% of eligible patients respectively. Exercise provocation had low feasibility with 38% of eligible patients unable to exercise due to factors other than dyspnea. A pDS was constructed from Likert scales during the three feasible assessment conditions (A–C). Relative to DVAS, the distribution of the pDS was more skewed with a high “ceiling effect” at enrollment (23%) limiting sensitivity to change. Change in pDS was not related to decongestion or 60-day outcomes.

Conclusions—While oxygen withdrawal and supine provocation are feasible and elicit more severe dyspnea, exercise provocation had unacceptable feasibility in this AHF cohort. The statistical characteristics of a pDS based on feasible provocation measures do not support its potential as a robust dyspnea assessment tool in AHF.

Clinical Trial Registration—RED-ROSE; [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01132846) identifier: NCT01132846

Keywords

Acute Heart Failure; Clinical Trials; Dyspnea

INTRODUCTION

Dyspnea relief is a primary goal of acute heart failure (AHF) therapy,(1, 2) a regulatory benchmark for the approval of novel therapeutic agents and a common endpoint in AHF clinical trials.(3–5) Most AHF trials have failed to demonstrate that tested interventions provided superior improvement in dyspnea as compared to placebo, potentially due to inadequate dyspnea assessment tools rather than ineffective therapeutic interventions.(5)

In most AHF trials, sequential dyspnea assessments using a categorical, 5-point-based Likert or continuous, 0 to 100-based visual analogue (DVAS) scale have been administered without standardization of conditions (oxygen use or position) which may modify dyspnea severity and influence change in scores over time. Moreover, seldom are patients enrolled on presentation to the emergency department when their dyspnea is most severe. Typically, enrollment and baseline assessments occur several hours after initiation of therapy at which point symptoms may have markedly improved. This further limits the sensitivity of conventional tools to detect an improvement in dyspnea. Supported by a study demonstrating that supine provocation elicited more severe dyspnea in AHF,(6) the AHF Syndromes International Working Group proposed a novel provocative dyspnea severity score (pDS) whereby dyspnea would be assessed under standardized, incrementally provocative maneuvers and called for studies to assess its feasibility and statistical characteristics.(7)

The Renal Optimization Strategies Evaluation in AHF (ROSE-AHF) trial compared the effect of placebo, low dose dopamine or low dose nesiritide on decongestion and renal function at 72 hours in patients with AHF. Dyspnea relief was assessed using DVAS without standardization of conditions as a secondary endpoint and did not vary by treatment group. (8) The Reliable Evaluation of Dyspnea in the ROSE study (RED-ROSE; [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01132846) identifier: NCT01132846) was an ancillary ROSE-AHF study designed to assess novel

symptom assessment tools in AHF. The objectives of this RED-ROSE analysis were to assess the feasibility and statistical characteristics of a pDS relative to the DVAS. As variable assessment conditions during DVAS administration could alter its relationship to the pDS, we first assessed the conditions (oxygen use and position) present during the DVAS assessments performed in ROSE-AHF. Second, we assessed the feasibility of provocative maneuvers and their impact on dyspnea severity as assessed by serial 5-point Likert scales. Finally, we assessed a fully constructed pDS in relation to the DVAS. While no quantitative physiologic measure of dyspnea severity exists to allow comparison of the predictive characteristics of the two scores against a gold standard, a more sensitive score should show a broader distribution and better correlate with measures of HF severity and treatment efficacy. Thus, we assessed the distribution of the two scores, their relationship to a marker of HF severity (NT-proBNP) on admission and their association with extent of decongestion or 60-day outcomes.

METHODS

Study Design

ROSE-AHF was performed within the National Heart Lung and Blood Institute (NHLBI)-sponsored HF Research Network (HFN) and enrolled 360 patients with AHF and renal dysfunction (estimated glomerular filtration rate (GFR) between 15 and 60 ml/min/1.73m²) within 24 hours of admission.(8) The diagnosis of AHF was based on at least one symptom (dyspnea, orthopnoea, or edema) and one sign of heart failure (rales, edema, ascites, or pulmonary vascular congestion on chest radiography) regardless of ejection fraction (EF). Electrolytes, serum NT-proBNP levels, body weight and symptom scores were assessed at enrollment (randomization), 24, 48 and 72-hours post-randomization. Cumulative urine output was assessed over 72 hours. Outcomes (hospitalization and vital status) were assessed at 60 days post-randomization.(9) RED-ROSE was approved by the NHLBI-appointed HFN data and safety monitoring board and by each site's institutional review board. Participants provided written informed consent.

Dyspnea VAS (DVAS)

At administration of the DVAS, patients were instructed to indicate how their breathing feels "right now" on an analogue scale from 0 (worst possible) to 100 (no breathlessness). In ROSE-AHF, the DVAS was assessed without pre-specified conditions but in RED-ROSE, the conditions (oxygen use, position) at DVAS assessment were recorded.

Feasibility and impact of provocative measures on dyspnea severity

The ROSE study coordinators were trained in the administration of the provocative measures and scales using scripts according to a standardized protocol. Subsequent to collection of the DVAS data, dyspnea severity was assessed with a 5-point Likert scale (worst possible, severe, moderate, mild or no dyspnea) during four stages [A. upright with oxygen (2 L/min), B. upright without oxygen, C. supine without oxygen, D. exercise without oxygen] which allowed assessment of changes in dyspnea severity with three provocative maneuvers (oxygen withdrawal, supine stress, exercise). Patients indicating moderate or less dyspnea on the Likert scale at each stage were eligible to proceed to the next stage. The

Likert scale was administered three minutes after starting stages A–C and immediately after the two minute exercise challenge (stage D). The exercise challenge consisted of stepping in place as quickly as possible at the bed side for two minutes. Prior to the step test, standing blood pressure was assessed. Patients were considered unable to exercise for non-dyspnea reasons if they refused to perform it or if they had lightheadedness, hypotension (systolic BP < 80 mmHg) on standing or mechanical limitations such as gait instability, arthritis or paralysis that were exercise-limiting. Patients who were on supplemental oxygen prior to enrollment were studied with the stipulation that oxygen levels be adjusted to 2 L/min during stage A and were only eligible to proceed to stage B (upright without oxygen) should they have reported moderate or less dyspnea as per Figure 2. Oxygen saturations were not assessed at any point.

Construction of the Provocative Dyspnea Severity Score (pDS)

The pDS proposed by AHF Syndromes International Working Group(7) outlined assessment of dyspnea with a Likert scale under five stages including upright with oxygen, upright without oxygen, supine without oxygen, after walking as fast as possible for 50 meters and after a six minute walk test (6MWT) and stipulated that patients with moderate or less severe dyspnea at each stage would proceed to the next. Due to concerns over feasibility of these types of exercise in newly admitted AHF patients across the range of inpatient settings involved in a multi-center trial, the simple bedside exercise provocation (stepping in place) described above was utilized. The proposed structure of the pDS incorporated the severity of dyspnea on the Likert scale, eligibility for different stages and the final stage tolerated.

Decongestion markers and clinical outcomes

Clinical markers of decongestion included weight change, cumulative urine volume and percent change in NT-proBNP from randomization to 72 hours. Clinical outcomes included all-cause death, HF rehospitalization or HF unscheduled outpatient visit (site investigator defined) at 60 days post-randomization.

Statistical Analysis

Patient characteristics were described as frequency (%) or median (interquartile range). Conditions at DVAS, feasibility of provocative measures, impact of provocative measures on dyspnea severity and score distributions were summarized using descriptive statistics. The Spearman correlation coefficient was used to examine the correlation between pDS and DVAS and their association with NT-proBNP at enrollment. The relationship between changes in scores and markers of clinical decongestion was examined using general linear models adjusting for baseline values of scores and baseline congestion markers. The association between changes in symptom scores and the time to an event was modeled using Cox Proportional Hazard regression models and Kaplan-Meier Curves. Wilcoxon rank test or Likelihood Ratio Chi-square tests were used to compare differences between groups. No imputation or carry forward was used to account for missing data. All analyses were conducted with SAS statistical software, version 9.2 or JMP, version 9.

This work was supported by grants from the National Heart, Lung, and Blood Institute (NHLBI) (U10 HL084904, U01 HL084861, U10 HL110312, U109 HL110337, U01

HL084889, U01 HL084890, U01 HL084891, U10 HL110342, U10 HL110262, U01 HL084931, U10 HL110297, U10 HL110302, U10 HL110309, U10 HL110336, U10 HL110338, HL 84907, T-91160, 5T32HL69749-10), the National Center for Advancing Translational Sciences (UL1TR000454, UL1TR000135, UL1RR025008, UL1TR000439) and the National Institute on Minority Health and Health Disparities (8 U54 MD007588). The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the paper and its final contents.

RESULTS

Baseline Characteristics

RED-ROSE commenced after ROSE-AHF had started and enrolled 232 of the ROSEAHF patients. Baseline characteristics of the RED-ROSE participants were similar to those of the ROSE cohort (Table 1). The median time from hospital admission to enrollment was 19.4 hours with an interquartile range of 15.6–21.9 hours. The median age was 69 years; 69% were men, and 25% were black (Table 1). A majority (62%) of patients had been hospitalized for HF within the previous 12 months. The median EF was 35% and 33% had an EF $\leq$ 50%. Cardiovascular comorbidities were common and 25% of patients had chronic obstructive pulmonary disease (COPD). On average, patients had stable hemodynamics, were obese, showed signs of volume overload, were anemic and had elevated NT-proBNP levels. Owing to the ROSE-AHF entry criteria, patients in RED-ROSE had moderate to severe renal dysfunction.

Conditions during DVAS assessment

DVAS data were available in 230 (99%) patients at enrollment (Supplemental Table 1). At the enrollment DVAS assessment, 88 (38%) patients were on oxygen and most assessments were performed with patients in an upright (64%) or semi-recumbent (28%) position and rarely in the supine (3%) or ambulatory (3%) position (position data missing in 2%). Oxygen use declined ($p < 0.0001$) but position distribution did not change over time (Supplemental Figure 1).

Feasibility of provocative maneuvers

Most patients (97%) were willing and able to perform the first three stages of the pDS (Supplemental Table 1). At enrollment, of the 160 patients eligible for the step test (Stage D), 61 (38%) did not perform it due to mechanical limitations ($n=23$), hypotension ($n=27$) or patient refusal ($n=11$) (Figure 1 and Supplemental Table 1). At 72 hours, 30% of eligible patients were still unwilling or unable to perform exercise for non-dyspnea reasons. Clinical characteristics of eligible patients who did or did not complete Stage D at enrollment were similar (Supplemental Table 2).

Effect of provocative maneuvers on dyspnea severity

At enrollment and 72 hours, oxygen withdrawal, supine position and activity each provoked more severe dyspnea (increase of one grade on a five point Likert dyspnea severity scale) in a significant proportion of patients (Figure 1).

pDS Construction

As a pDS must have high feasibility to be used as a clinical trial endpoint, the pDS was a nine-point score constructed from stages (A–C) which had acceptable feasibility (Figure 2).

Distribution of DVAS and pDS

At enrollment, the distribution of pDS was skewed (Figure 3 A) as the majority of patients did not have “worst possible” or “severe” dyspnea during Stages A and B and thus, were scored according to their dyspnea severity in Stage C. In contrast, patients were more evenly distributed across the range of possible DVAS scores (Figure 3 B). The ceiling effect (% of patients with highest possible score) was higher for pDS (23% at enrollment, 43% at 72 hours) than for DVAS (3.0% at enrollment, 7.9% at 72 hours). At enrollment and 72 hours, pDS and DVAS were correlated ($p < 0.001$) but DVAS scores varied widely within each pDS score (Figure 4).

Correlation of DVAS and pDS with HF severity

At enrollment, neither DVAS nor pDS were correlated with the NT-proBNP level ($|r| < 0.03$, $p > 0.6$ for both).

Change in dyspnea scores over time

The pDS and DVAS each increased over time indicating dyspnea improvement ($p < 0.0001$ for each; Figure 3 A and B).

Change in weight and net fluid loss at 72 hours were correlated ($r = 0.57$, $p < 0.001$) but neither weight change nor fluid loss were correlated with percent change in NT-proBNP ($|r| < 0.03$, $p > 0.05$ for both). There were statistically significant associations between changes in DVAS and pDS from baseline to 72 hours and extent of decongestion at 72 hours but these were not clinically meaningful as the model R^2 values were all quite low (< 0.05 ; Table 2). There were no significant associations between changes in DVAS or pDS from baseline to 72 hours and 60-day outcomes (Log Rank $p > 0.70$ for all; Supplemental Figure 2).

There were also no clinically meaningful relationships between either score at 72 hours and decongestion markers or outcomes (data not shown).

DISCUSSION

In this exploratory study, we examined the feasibility and statistical characteristics of a novel dyspnea assessment tool (pDS) in relation to a standard tool (DVAS) widely used in AHF trials. Provocative maneuvers each elicited more severe dyspnea in a significant portion of patients but exercise stress was not feasible enough to employ as part of an endpoint in an AHF clinical trial. The distribution of a pDS constructed with feasible provocation maneuvers was more highly skewed than that of the DVAS, with a high ceiling effect limiting sensitivity to change. Relative to DVAS, the pDS was not more closely related to a biomarker of HF severity (NT-proBNP), markers of decongestion or 60-day outcomes. Our findings do not support utility of a pDS in AHF trials.

Dyspnea assessment tools in AHF

The most widely used measures of dyspnea in AHF trials are the DVAS and the Likert scale. (7) Most AHF trials have not stipulated standardized conditions during dyspnea score administration. (7) We found that oxygen use was common at enrollment DVAS assessment and declined over time and that few patients were supine during assessments. Further, characterization of dyspnea on a visual analogue scale did not appear to be tightly correlated with Likert based, ordinal dyspnea descriptors as DVAS varied widely (Figure 4) across descriptor terms in the pDS. These findings may contribute to variable and on average, modest changes in DVAS scores over time in RED-ROSE and other studies. (10, 11)

Feasibility of provocative measures in AHF

In RED-ROSE, exercise provocation was feasible in only 62% of eligible patients at enrollment. Other studies reporting feasibility of exercise in AHF have excluded patients with non-HF conditions affecting mobility (12–14) and generally enrolled younger patients with less severe renal dysfunction (13, 14). A study without exclusion for limited mobility and with a mean age similar to RED-ROSE, reported that 69% of patients could perform a walk test at discharge, (15) similar to our findings where 70% of eligible patients could perform the step test at 72 hours after enrollment. Thus, while exercise provocation did induce more severe dyspnea in 29% of patients, feasibility was too low to allow incorporation into a clinical trial endpoint.

Impact of provocative measures on dyspnea severity in AHF

Owing to the design of ROSE/RED-ROSE and similar to those of other contemporary AHF trials, (10, 11, 16) patients could be enrolled up to 24 hours after admission to the hospital. The minority of AHF trial patients are enrolled on presentation to the emergency department when their dyspnea is most severe. Typically, enrollment and baseline assessments occur several hours after initiation of therapy at which time resting dyspnea may have markedly improved. This supports the case for provocative measures to better expose dyspnea severity.

Only one previous study, the URGENT-dyspnea study, used a provocative dyspnea assessment tool in AHF but limited provocation to supine stress performed only on admission and showed that resting dyspnea improves rapidly in the first 6 hours after presentation. (6) The findings of RED-ROSE extend this study using a more detailed provocation protocol and serial assessments. Moreover, despite enrollment occurring much later during the hospitalization in RED-ROSE compared to URGENT-dyspnea (within 24 versus 1 hour of presentation), postural stress induced more severe dyspnea in 42% of patients, similar to that observed in the URGENT-dyspnea study. Oxygen withdrawal and exercise (not assessed in the URGENT-dyspnea study) induced more severe dyspnea in 24% of patients. These findings suggest that regardless of enrollment time, conditions at dyspnea assessments influence dyspnea severity and should be considered in AHF trial design.

Statistical characteristics and clinical correlates of the pDS

The skewedness of the pDS, the high ceiling effect at enrollment and the fact that exercise did induce more severe dyspnea in a third of the patients able to perform it suggest that

additional provocation measures beyond supine stress would be needed to provide a dyspnea assessment tool with high sensitivity to dyspnea severity, particularly milder dyspnea. Thus, the limited feasibility of exercise in the RED ROSE trial is disappointing.

Alternate forms of activity provocation (hand grip, supine cycle ergometry) which could be performed in bed could be considered but would add complexity and require special equipment in an AHF trial. Alternatively, the skewed pDS distribution may reflect the fact that resting dyspnea is not always present in AHF patients (17) and even when present, may have improved between presentation and enrollment in ROSE-AHF at up to 24 hours after presentation. Indeed, the majority of patients who performed the step test did not report more severe dyspnea than noted during the supine assessment, suggesting that a substantial portion of patients had little capacity for improvement in dyspnea.

Changes in dyspnea scores at 72 hours were not meaningfully correlated with changes in markers of decongestion or with 60-day outcomes, consistent with some(18) but not all previous studies(19, 20) in AHF. While power to detect association of changes in dyspnea with decongestion or outcomes in RED-ROSE was limited by sample size, our findings do not provide evidence that the pDS is more sensitive to treatment efficacy than DVAS.

The limited feasibility of exercise provocation in AHF, the fact that conditions for DVAS assessment varied widely and changed over time when not pre-specified, and the finding that feasible maneuvers such as oxygen withdrawal and supine position provoke more severe dyspnea suggest that simply standardizing resting conditions (supine without oxygen) for serial VAS or Likert based dyspnea assessments would be highly feasible and likely to enhance the sensitivity to change in dyspnea over time in AHF trials. One completed (11) and some ongoing AHF trials (NCT01966601 and NCT01300013) are employing such standardization and emphasizing enrollment earlier after presentation to capture early improvement.

Limitations

Of all the patients enrolled in RED-ROSE, 28% had COPD; thus, the effect of the stage A→B transition of the pDS may be a reflection of their COPD and not a feature likely to improve with decongestion. A more demanding physical activity performed in stage D of the pDS such as a 6MWT or stair climbing rather than bedside stepping in place may have provoked more severe dyspnea and potentially allowed room to demonstrate improvement in pDS with AHF treatment. Nonetheless, such maneuvers may not be feasible in an AHF population. Neither fluid output nor weight change was scaled to an assessment of the goal fluid or weight loss. Thus, these widely used measures are only rough indices of decongestion efficacy. Exclusion of AHF patients without significant renal dysfunction may have affected our findings. A VAS-based pDS may have yielded different findings. Findings may be different in studies enrolling patients closer to initial presentation. Females were underrepresented in this cohort.

Conclusion

As oxygen use and position at dyspnea assessments vary widely in the absence of pre-specified assessment conditions and affect dyspnea severity regardless of enrollment time,

serial dyspnea assessments should be made under standardized resting conditions (supine without oxygen) in AHF clinical trials. While oxygen withdrawal and supine stress are feasible and provoke more severe dyspnea, exercise stress had unacceptable feasibility in this AHF study. The statistical characteristics of a pDS based on feasible provocation measures do not support its potential as a robust dyspnea assessment tool in AHF. Our findings highlight the importance for future AHF trials to assess symptoms in a standardized fashion and as early as possible during an AHF admission in order to increase the sensitivity of conventional tools to detect an improvement in dyspnea and better classify AHF symptoms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING

This work was supported by grants from the National Heart, Lung, and Blood Institute (NHLBI) (coordinating center: U01 HL084904; regional clinical centers and Heart Failure Network Clinical Research Skills Development Cores : U01 HL084861, U01 HL110312, U01 HL110337, U01 HL084889, U01 HL084890, U01 HL084891, U01 HL110342, U01 HL110262, U01 HL084931, U01 HL110297, U01 HL110302, U01 HL110309, U01 HL110336, U01 HL110338). This work was also supported by the National Center for Advancing Translational Sciences (UL1TR000454, UL1TR000135, UL1RR025008, UL1TR000439) and the National Institute on Minority Health and Health Disparities (8 U54 MD007588). Dr. Omar AbouEzzeddine was funded by the NIH Training Grant (T-91160) and the HFCRN Skills Development Core (HL 84907). Dr. Prateeti P. Khazanie was supported by grant 5T32HL69749-10 from the National Heart, Lung, and Blood Institute.

Dr Chen is the co-founder of Zumbro Discovery. Consultant for: Bayer, BG Medicine, the Medicines Company, Otsuka, Intersection Medical, Janssen, Medtronic, Novartis, Trevena, scPharmaceuticals, Cardioxyl, Roche Diagnostics. Honoraria: Relypsa, Palatin Technologies, Abbott Research Grants: NHLBI/NCATS, Alere. **AFH:** However, he reports grants from Amgen, Bristol Myers Squibb, GlaxoSmithKline and Novartis, outside the submitted work.

REFERENCES

1. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2013; 62:e147–e239. [PubMed: 23747642]
2. Lindenfeld J, Albert NM, Boehmer JP, Collins SP, Ezekowitz JA, Givertz MM, Katz SD, Klapholz M, Moser DK, Rogers JG, Starling RC, Stevenson WG, Tang WH, Teerlink JR, Walsh MN. HFSA 2010 Comprehensive Heart Failure Practice Guideline. *J Card Fail.* 2010; 16:e1–e194. [PubMed: 20610207]
3. West RL, Hernandez AF, O'Connor CM, Starling RC, Califf RM. A review of dyspnea in acute heart failure syndromes. *Am Heart J.* 2010; 160:209–214. [PubMed: 20691823]
4. Teerlink JR. Dyspnea as an end point in clinical trials of therapies for acute decompensated heart failure. *Am Heart J.* 2003; 145:S26–S33. [PubMed: 12594449]
5. Pang PS, Komajda M, Gheorghiade M. The current and future management of acute heart failure syndromes. *Eur Heart J.* 2010; 31:784–793. [PubMed: 20207624]
6. Mebazaa A, Pang PS, Tavares M, Collins SP, Storrow AB, Laribi S, Andre S, Mark Courtney D, Hasa J, Spinar J, Masip J, Frank Peacock W, Sliwa K, Gayat E, Filippatos G, Cleland JG,

Gheorghiade M. The impact of early standard therapy on dyspnea in patients with acute heart failure: the URGENT-dyspnea study. *Eur Heart J*. 2010; 31:832–841. [PubMed: 19906690]

7. Pang PS, Cleland JG, Teerlink JR, Collins SP, Lindsell CJ, Sopko G, Peacock WF, Fonarow GC, Aldeen AZ, Kirk JD, Storrow AB, Tavares M, Mebazaa A, Roland E, Massie BM, Maisel AS, Komajda M, Filippatos G, Gheorghiade M. A proposal to standardize dyspnea measurement in clinical trials of acute heart failure syndromes: the need for a uniform approach. *Eur Heart J*. 2008; 29:816–824. [PubMed: 18310669]
8. Chen HH, Anstrom KJ, Givertz MM, Stevenson LW, Semigran MJ, Goldsmith SR, Bart BA, Bull DA, Stehlik J, LeWinter MM, Konstam MA, Huggins GS, Rouleau JL, O'Meara E, Tang WH, Starling RC, Butler J, Deswal A, Felker GM, O'Connor CM, Bonita RE, Margulies KB, Cappola TP, Ofili EO, Mann DL, Dávila-Román VG, McNulty SE, Borlaug BA, Velazquez EJ, Lee KL, Shah MR, Hernandez AF, Braunwald E, Redfield MM. Low-dose dopamine or low-dose nesiritide in acute heart failure with renal dysfunction: the ROSE acute heart failure randomized trial. *JAMA*. 2013; 310:2533–2543. [PubMed: 24247300]
9. Chen HH, AbouEzzeddine OF, Anstrom KJ, Givertz MM, Bart BA, Felker GM, Hernandez AF, Lee KL, Braunwald E, Redfield MM. Targeting the kidney in acute heart failure: can old drugs provide new benefit? Renal Optimization Strategies Evaluation in Acute Heart Failure (ROSE AHF) trial. *Circ Heart Fail*. 2013; 6:1087–1094. [PubMed: 24046475]
10. McMurray JJ, Teerlink JR, Cotter G, Bourge RC, Cleland JG, Jondeau G, Krum H, Metra M, O'Connor CM, Parker JD, Torre-Amione G, van Veldhuisen DJ, Lewsey J, Frey A, Rainisio M, Kobrin I. Effects of tezosentan on symptoms and clinical outcomes in patients with acute heart failure: the VERITAS randomized controlled trials. *JAMA*. 2007; 298:2009–2019. [PubMed: 17986694]
11. Teerlink JR, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Ponikowski P, Unemori E, Voors AA, Adams KF Jr, Dorobantu MI, Grinfeld LR, Jondeau G, Marmor A, Masip J, Pang PS, Werdan K, Teichman SL, Trapani A, Bush CA, Saini R, Schumacher C, Severin TM, Metra M. Serelaxin, recombinant human relaxin-2, for treatment of acute heart failure (RELAX-AHF): a randomised, placebo-controlled trial. *Lancet*. 2013; 381:29–39. [PubMed: 23141816]
12. Kommuri NV, Johnson ML, Koelling TM. Six-minute walk distance predicts 30-day readmission in hospitalized heart failure patients. *Arch Med Res*. 2010; 4:363–368. [PubMed: 20851294]
13. Alahdab MT, Mansour IN, Napan S, Stamos TD. Six minute walk test predicts long-term all-cause mortality and heart failure rehospitalization in African-American patients hospitalized with acute decompensated heart failure. *J Card Fail*. 2009; 15:130–135. [PubMed: 19254672]
14. Howie-Esquivel J, Dracup K. Effect of gender, ethnicity, pulmonary disease, and symptom stability on rehospitalization in patients with heart failure. *Am J Cardiol*. 2007; 100:1139–1144. [PubMed: 17884378]
15. Feola M, Garnero S, Vallauri P, Salvatico L, Vado A, Leto L, Testa M. Relationship between cognitive function, depression/anxiety and functional parameters in patients admitted for congestive heart failure. *Open Cardiovasc Med J*. 2013; 7:54–60. [PubMed: 24044027]
16. Felker GM, Lee KL, Bull DA, Redfield MM, Stevenson LW, Goldsmith SR, LeWinter MM, Deswal A, Rouleau JL, Ofili EO, Anstrom KJ, Hernandez AF, McNulty SE, Velazquez EJ, Kfoury AG, Chen HH, Givertz MM, Semigran MJ, Bart BA, Mascette AM, Braunwald E, O'Connor CM, NHLBI Heart Failure Clinical Research Network. Diuretic strategies in patients with acute decompensated heart failure. *N Engl J Med*. 2011; 364(9):797–805. [PubMed: 21366472]
17. Kato M, Stevenson LW, Palardy M, Campbell PM, May CW, Lakdawala NK, Stewart G, Nohria A, Rogers JG, Heywood JT, Gheorghiade M, Lewis EF, Mi X, Setoguchi S. The worst symptom as defined by patients during heart failure hospitalization: implications for response to therapy. *J Card Fail*. 2012; 18:524–533. [PubMed: 22748485]
18. Kociol RD, McNulty SE, Hernandez AF, Lee KL, Redfield MM, Tracy RP, Braunwald E, O'Connor CM, Felker GM. Markers of decongestion, dyspnea relief, and clinical outcomes among patients hospitalized with acute heart failure. *Circ Heart Fail*. 2013; 6:240–245. [PubMed: 23250981]
19. Mentz RJ, Hernandez AF, Stebbins A, Ezekowitz JA, Felker GM, Heizer GM, Atar D, Teerlink JR, Califf RM, Massie BM, Hasselblad V, Starling RC, O'Connor CM, Ponikowski P. Predictors

- of early dyspnea relief in acute heart failure and the association with 30-day outcomes: findings from ASCEND-HF. *Eur J Heart Fail.* 2013; 15:456–464. [PubMed: 23159547]
20. Metra M, O'Connor CM, Davison BA, Cleland JG, Ponikowski P, Teerlink JR, Voors AA, Givertz MM, Mansoor GA, Bloomfield DM, Jia G, DeLucca P, Massie B, Dittrich H, Cotter G. Early dyspnea relief in acute heart failure: prevalence, association with mortality, and effect of rolofylline in the PROTECT Study. *Eur Heart J.* 2011; 32:1519–1534. [PubMed: 21388992]

Figure 1. Impact of provocative measures on dyspnea severity at enrollment and 72 hours
 The percent of patients who had less severe (“better”, improvement by 1 point on Likert dyspnea scale), “unchanged”, or “worse” (worsening by 1 point on Likert dyspnea scale) dyspnea or who were eligible for but did perform the stage due to mechanical limitations, hypotension or refusal (unable) is shown for each provocative maneuver (oxygen withdrawal (O₂ WD), supine position or bedside step test).

Figure 2. Provocative dyspnea score (pDS)

The proposed structure of the pDS was a nine-point score that incorporated 1) the severity of dyspnea on the Likert scale at three progressively more provocative stages (A: upright with O₂, B: upright without O₂, C: supine), 2) the eligibility for different stages and 3) the final stage tolerated.

Figure 3. Distribution of dyspnea scores at enrollment

The frequency distribution of the provocative dyspnea severity score (pDS, A) and the dyspnea visual analogue scale (DVAS, B) for patients with paired pDS and DVAS both at enrollment (“baseline”) and 72 hours (n=203) are shown. Distributions were similar when all patients with pDS or DVAS at baseline or 72 hours were included.

Figure 4. Distribution of the dyspnea visual analogue scale (DVAS) score relative to the provocative dyspnea score (pDS)

Scatter dot plot of DVAS with medians (red bar) according to the pDS scores for patients with paired pDS and DVAS both at baseline (A) and 72 hours (B)(n=203) are shown. DVAS increased with increasing pDS ($p < 0.0001$ for both) but varied widely within each pDS score. Distributions of DVAS relative to the pDS were similar when all patients paired pDS and DVAS at baseline or 72 hours were included.

Table 1**Baseline Characteristics**

Characteristic	RED ROSE (N=232)	ROSE (N=360)
Age, years	69 (62–79)	70 (62–79)
Male sex	159 (69%)	264 (73%)
White race	166 (72%)	272 (76%)
HF hospitalization	143 (62%)	240 (67%)
Ejection fraction, %	35 (23–54)	34 (21–53)
Ejection fraction $\geq 50\%$	76 (33%)	94 (26%)
Hypertension	195 (84%)	298 (83%)
Diabetes	127 (55%)	200 (56%)
Stroke	21 (9%)	31 (9%)
Atrial fibrillation	135 (58%)	215 (60%)
COPD	59 (25%)	95 (26%)
Medications		
Loop diuretic	215 (93%)	340 (94%)
ACE or ARB	112 (48%)	179 (50%)
Beta blocker	196 (84%)	300 (83%)
Aldosterone antagonist	68 (29%)	109 (30%)
Digoxin	53 (23%)	89 (25%)
Clinical Examination		
Heart rate (bpm)	74 (66–85)	74 (65–85)
Systolic BP, mmHg	116 (104–129)	114 (103–127)
Body mass index, Kg/m ²	31 (27–37)	31 (27–37)
JVP ≥ 8 cm	212 (95%)	327 (95%)
Rales	122 (54%)	197 (56%)
Edema $\geq 2+/4+$	159 (69%)	251 (70%)
Orthopnoea	193 (88%)	307 (90%)
Laboratory data		
Haemoglobin, g/dL	11.5 (10.4–12.8)	11.4 (10.3–12.7)
eGFR, mL/min/1.73m ²	45 (33–56)	42 (32–53)
NT Pro-BNP, pg/mL	5055 (2358–10348)	5323 (2420–10797)
Baseline dyspnea scores		
DVAS (Scale 0–100)	62.5 (40.0 – 82.0)	60.0 (39.0 – 80.0)
pDS (Scale 1–9)	7 (6–8)	–

Data are number (%) or median (interquartile range).

Abbreviations:

ACE, Angiotensin Converting Enzyme; ARB, Angiotensin Receptor Blocker; BP, Blood Pressure; COPD, Chronic Obstructive Pulmonary Disease; eGFR, Estimated Glomerular Filtration Rate; HF, Heart Failure; JVP, Jugular Venous Pressure; NT Pro-BNP, N-terminal pro-brain natriuretic peptide

Table 2

Association between changes in symptom scores and extent of decongestion.

	Cumulative Urine Volume (BL to 72 Hours)			Weight Change (BL to 72 Hours)			Percent Change in NT-proBNP (BL to 72 Hours)		
	Model R ²	Parameter Estimate (CI) ml	P value	Model R ²	Parameter Estimate (CI) lbs	P value	Model R ²	Parameter Estimate (CI) %	P value
Change DVAS (BL to 72 Hrs)	0.04	29 (9 – 50)	0.005	0.02	-0.05 (-0.10 – 0.004)	0.07	0.01	-0.23 (-0.50 – 0.04)	0.10
Change pDS (BL to 72 Hrs)	0.01	233 (-116 – 583)	0.19	0.04	-1.15 (-2.01 – -0.28)	0.01	0.01	-2.88 (-7.26 – 1.51)	0.20

Model adjusts for baseline symptom score value and when appropriate, for baseline value of the congestion marker (weight and NT-proBNP level).